

Valeria Ottonelli

Il previsore di Newcomb: un paradosso o un problema mal posto?

Abstract - Il paradosso di Newcomb ha goduto per anni di una notevole popolarità fra gli studiosi di lingua inglese, ed ha suscitato innumerevoli ed interminabili discussioni, senza che per questo si sia mai arrivati ad una sua soluzione o ad un consenso fra i diversi punti di vista che sono sorti al suo riguardo.

Alcuni studiosi hanno cercato di spiegare il perché della irriducibilità delle tesi opposte che sono sorte dalla discussione del problema di Newcomb. Queste spiegazioni hanno fatto riferimento alla decision theory e alla teoria dei giochi ed hanno tentato di ridurre la discussione sul paradosso ad una discussione sui principi di scelta in situazioni di rischio.

Le spiegazioni avanzate sino a questo momento non sembrano soddisfacenti perché appaiono viziate da presupposti teorici discutibili e perché non sembrano affrontare il problema alla radice e in tutta la sua complessità.

E' possibile mostrare che in realtà il problema di Newcomb è un problema mal posto, e che il vero motivo per cui non si può arrivare ad una soluzione pienamente accettabile è che la sua formulazione è sufficientemente ambigua da lasciare sempre spazio per un disaccordo sulla sua reale natura.

La formulazione classica del problema di Newcomb contiene almeno tre punti oscuri, o passibili di interpretazioni contraddittorie. Dall'esame dettagliato di tali punti emerge chiaramente la soluzione delle maggiori controversie originate dalla discussione del problema. Ciò che emerge, inoltre, è che, se si chiariscono le ambiguità contenute nella formulazione del problema di Newcomb, ci si trova di fronte non ad uno, ma ad almeno tre problemi diversi, e di semplice soluzione. Le soluzioni a questi tre semplici problemi corrispondono ai tre principali tipi di argomentazione usati da quanti hanno discusso il paradosso nella sua formulazione originaria.

Questa discussione del problema dovrebbe mostrare:

- 1) che il paradosso di Newcomb è un problema mal posto
- 2) che a tale problema teorico non si può dare una vera e propria soluzione
- 3) che ciascuno dei tre principali tipi di argomentazione proposti da quanti hanno discusso il problema ha una sua plausibilità, che non può essere messa in discussione
- 4) che tuttavia tutti e tre i tipi di argomentazione sono fondati su delle fallacie, non imputabili ai loro autori ma alla formulazione del problema stesso

Nelle pagine che seguiranno intendo discutere un paradosso che gode da ormai un ventennio di una certa popolarità, e che viene solitamente designato col nome del suo creatore, il fisico William Newcomb. A partire dalla sua formulazione classica sono state elaborate innumerevoli varianti, ma ritengo che si possa individuare una definizione standard che ne riassume i termini essenziali e che potrebbe essere formulata nel seguente modo:

mi trovo di fronte a due scatole, la scatola n° 1 e la scatola n° 2, e mi sono proposte due scelte possibili: posso prendere entrambe le scatole oppure prendere solo la scatola n°1. Mi viene inoltre fornita la seguente informazione, che io so essere corretta: qualsiasi cosa io scelga, ci sono mille dollari nella scatola n°2, ma esiste un essere dotato di straordinari poteri di previsione, nei quali io ho

un'enorme fiducia, che ha previsto la scelta che farò, e ha messo un milione di dollari nella scatola n°1 se e solo se ha previsto che io avrei preso solo quest'ultima. Che cosa è razionale che io scelga in questa situazione?

Il mio intento è quello di dimostrare che il dilemma appena presentato non è un problema ben posto, e che gran parte delle discussioni che ne sono state originate hanno in realtà abusato delle ambiguità contenute nella sua formulazione.

Negli ultimi due decenni il problema di Newcomb ha goduto di una notevole popolarità fra gli studiosi di lingua inglese. La letteratura relativa è quindi molto vasta, e sarebbe estremamente difficile, oltretutto fuori luogo, fornire una ricostruzione completa ed accurata di tutte le discussioni del problema che sono state proposte in questi anni. Tuttavia, sarà utile fornire una descrizione approssimativa e semplificata dei principali argomenti a favore di ciascuna delle due soluzioni possibili del problema, cioè degli argomenti che conducono alla scelta di entrambe le scatole, e di quelli che conducono alla scelta di una scatola sola:

1) chi argomenta a favore della scelta di una scatola sola sostiene che, sulla base delle statistiche, si è indotti a credere ad un'alta percentuale di riuscita delle previsioni. Ciò significa che, almeno fino al momento della scelta, e nella maggior parte dei casi, è risultato che chi ha preso entrambe le scatole ha guadagnato solo un milione, mentre chi ha preso solo una scatola ha guadagnato un miliardo. Quindi, statisticamente, la scelta di una scatola sola si è rivelata la più felice. Esiste inoltre una seconda argomentazione a favore di questa scelta: poiché il previsore è quasi perfetto, se si sceglie una sola scatola ne conseguirà quasi necessariamente (o molto probabilmente) che il previsore lo avrà previsto, e che quindi si guadagnerà un miliardo, mentre se si scelgono entrambe le scatole, ne conseguirà altrettanto probabilmente che il previsore lo avrà previsto, e che quindi si guadagnerà solo un milione.

2) i sostenitori della scelta di entrambe le scatole giustificano tale decisione nel seguente modo: qualsiasi azione uno decida di fare, il previsore avrà già messo i soldi nelle scatole, e quindi il loro contenuto non potrà subire alcuna variazione. Quale che sia il contenuto, conviene ovviamente prendere entrambe le scatole piuttosto che una sola.

Due questioni fondamentali

Il problema principale che deve essere risolto da chi affronta questo paradosso non è ovviamente quello di stabilire quale sia la decisione corretta, ma il problema di spiegare perché è possibile un disaccordo su tale questione. In

effetti, una delle caratteristiche più interessanti del paradosso di Newcomb sembra essere proprio il fatto di generare fazioni opposte ed irriducibili, senza che apparentemente intervengano fattori che siano in grado di determinare una contrapposizione di questo genere, come succederebbe invece se ad esempio fosse possibile introdurre premesse valutative diverse. Nella descrizione del problema sembrano comparire tutti i dati rilevanti per la soluzione, senza che nulla venga lasciato alla fantasia di chi deve scegliere. Come è possibile che da questa situazione emergano punti di vista contrapposti?

Esiste inoltre un altro fenomeno a prima vista inspiegabile: nonostante l'apparente irriducibilità delle due tesi opposte, esistono alcune condizioni alle quali sembra esserci una convergenza fra le due fazioni. Nel caso in cui l'evidenza statistica a favore del previsore raggiungesse la misura del 100%, alcuni dei sostenitori della tesi delle due scatole sarebbero disposti a cambiare idea, e ad entrare a far parte della schiera contrapposta. In una situazione di questo genere, infatti, l'argomento a favore di una scatola sola sembrerebbe acquisire una particolare forza, poiché la connessione fra la scelta e l'esito apparirebbe necessaria. Il problema, in realtà, lascia del tutto imprecisata la misura dell'evidenza probabilistica; ma, proprio grazie a questa indeterminatezza, rimane possibile l'esperimento mentale di inserire fra i dati del problema misure sempre più prossime al 100%, e assistere così allo strano fenomeno per cui, allo scattare della percentuale piena, i sostenitori della tesi delle due scatole si trasformano in sostenitori della tesi opposta.

Le spiegazioni proposte per rispondere al primo problema sono state avanzate principalmente dai sostenitori della tesi a favore di entrambe le scatole. Il tipo di argomentazioni usate da questi ultimi fa riferimento alla *decision theory* e agli strumenti concettuali da questa elaborati. La tesi avanzata può essere riassunta nel seguente modo: esistono due fondamentali principi di scelta in situazioni di rischio: il *principio di dominanza* ed il *principio di massimizzazione dell'utilità prevista*. Il primo impone di scegliere, fra due azioni disponibili, quella "dominante", cioè quella il cui esito è preferito all'esito dell'alternativa in almeno uno degli scenari possibili (nel caso del problema di Newcomb gli scenari possibili sono due: quello in cui la scatola n° 1 contiene il milione di dollari, e quello in cui non lo contiene), mentre in nessuno di essi il suo esito è peggiore di quello dell'altra. Il secondo principio asserisce invece che l'azione migliore disponibile è quella che massimizza l'utilità prevista, cioè il valore della sommatoria dei prodotti della probabilità degli esiti possibili dell'azione e dei loro valori di utilità per l'agente. Nella maggior parte dei casi l'applicazione di questi due principi dà luogo agli stessi risultati. Ma esistono delle situazioni in cui ciò non è vero, ed il problema di Newcomb è una di queste. Tra i casi in cui i due principi entrano in conflitto (danno risposte discordanti) esistono quelli in cui le decisioni non hanno influenza sulle condizioni di scelta (benché possano

averne sugli esiti): nel caso particolare del problema di Newcomb, ad esempio, la scelta può influire sul proprio esito (a seconda che si prendano entrambe le scatole o se ne prenda una sola cambia l'ammontare del guadagno), ma non può influenzare le proprie condizioni di partenza (qualsiasi cosa si scelga, la disposizione dei soldi nelle scatole non cambia). In questi casi, secondo i sostenitori della tesi a favore della scelta di entrambe le scatole, a prevalere dev'essere il principio di dominanza, perché il principio dell'utilità prevista non tiene conto della differenza fra le situazioni in cui le condizioni di partenza sono modificabili e quelle in cui non lo sono, e si comporta come se l'agente fosse sempre in grado di influire sul contesto di scelta. Secondo i sostenitori di questa tesi, quindi, chi sceglie la seconda scatola commette l'errore di applicare un principio di scelta inadeguato alla situazione. Questa spiegazione vale ovviamente al tempo stesso anche come giustificazione della tesi per cui bisogna scegliere entrambe le scatole.

La spiegazione appena presentata è inficiata da un grave errore epistemologico. Essa presuppone che nel problema di Newcomb non esista una relazione causale diretta fra ciò che si sceglie ed il contenuto delle scatole, e che i risultati statistici siano spiegabili con l'esistenza di una causa comune ai due fenomeni. Molti dei sostenitori della tesi delle due scatole sono fertili creatori di esempi apparentemente analoghi alla situazione descritta dal problema di Newcomb. Questi esempi dovrebbero servire a mostrare come la co-variazione statistica fra due fenomeni diversi non sia sempre sufficiente a stabilire l'esistenza di rapporti causali fra i fatti in questione, poiché in molti casi è possibile indicare la presenza di un terzo fatto che, essendo causa di entrambi, spieghi la loro sintonia. Tali esempi potrebbero essere divisi in due classi: quelli che includono fra i dati tutta l'evidenza comunemente ritenuta necessaria per stabilire l'esistenza di una causa di questo genere, e che quindi non hanno nulla a che fare col problema di Newcomb, e quelli che non includono fra i dati tale evidenza, e che quindi sono sottoponibili alle stesse obiezioni cui va incontro il paradosso di Newcomb. Questa seconda classe di esempi include tutti quelli in cui il fatto che verrebbe istintivamente proposto come causa di un fenomeno da spiegare, e che viene poi successivamente eliminato dalla descrizione della relazione causale, è un'azione o una decisione umana. Poiché nessuna scienza è mai arrivata alla descrizione di rapporti causali precisi fra fatti e decisioni umane, mentre è possibile appurare l'esistenza di rapporti precisi fra azioni umane e conseguenze fattuali, gli esempi di questo genere sono generalmente implausibili. Di solito, tutto quello che i loro creatori riescono a fare è inventare una caratteristica umana sufficientemente ambigua e generica da poter essere indicata come causa di particolari azioni, senza che la relazione così delineata possa venire messa a confronto coi fatti (la causa di una particolare decisione è la "disposizione a decidere in quel modo"). In nessun caso essi riescono a descrivere

situazioni convincenti, in cui risulti chiaro che la descrizione corretta è quella che non mette in relazione l'azione umana in questione con un certo esito. L'unico modo di confutare la spiegazione di una covariazione statistica fra fatti è presentare *altre statistiche*, e non altre spiegazioni che fanno ricorso a fatti inesistenti o non suscettibili di prova. L'errore non solo concettuale che si deve imputare ai creatori di questi esempi è quello di immaginare che una discussione filosofica sia in grado di scoprire in una data situazione nessi causali che altrimenti sfuggirebbero.

I sostenitori di questo tipo di teorie fingono quindi di dimostrare ciò che hanno postulato. E' vero che l'uso scorretto di alcuni principi può mascherare la vera natura del contesto di scelta, ma non è altrettanto vero che la scelta dei principi debba, una volta stabilito questo, essere pacifica. Da queste stesse discussioni emerge chiaramente che il problema reale è quello della descrizione del contesto di scelta. Lo stesso principio, se applicato a situazioni differenti (o meglio: a descrizioni differenti della stessa situazione), produce esiti differenti. Allora la discussione dovrebbe concentrarsi principalmente sulle descrizioni che stanno alla base dell'applicazione dei principi. Nel caso di Newcomb buona parte delle controversie riguarda infatti la possibilità di una relazione causale che proceda a ritroso attraverso il tempo, cioè la possibilità di una descrizione della situazione che è completamente diversa da quella assunta dai sostenitori della tesi delle due scatole. Non è corretto, allora, attribuire la divergenza delle descrizioni della rete di relazioni causali sottostante al problema di Newcomb alla diversa formulazione dei principi di scelta, dal momento che questa è evidentemente una conseguenza, e non una causa, di tale divergenza delle descrizioni. Il problema è quindi quello di capire come sia possibile che qualcuno assuma come corretta la descrizione che conduce alla scelta di una sola scatola, e quale sia il genere di argomenti che sostiene la plausibilità di questa descrizione.

Alla seconda domanda, quella sul motivo per cui un successo passato del 100% da parte del previsore genera concordia fra le due fazioni, non è stata data alcuna risposta convincente. I sostenitori della tesi delle due scatole che prendono in considerazione il problema tendono semplicemente a giustificarsi affermando che il 100% di successi crea una situazione diversa da quella originaria.

Nelle pagine che seguiranno tenterò di dare una risposta ad entrambe le questioni. Purtroppo tale risposta non sarà anche contemporaneamente una soluzione al problema di Newcomb, e questo semplicemente perché penso che il problema non possa avere una vera e propria risposta corretta. Cercherò invece di spiegare perché coloro che si sono occupati del problema hanno fornito soluzioni così contrastanti ed irriducibili, e come questo dipenda dalla natura del problema stesso. I bersagli principali della mia discussione non saranno dunque i vari tentativi di giustificare particolari scelte di fronte al problema di Newcomb,

ma le teorie che fanno riferimento ai principi di scelta e alle loro condizioni, che tentano cioè di spiegare le differenze di vedute con la presenza di errori intellettuali nell'uso dei principi di scelta.

Le trappole di newcomb

Il motivo per cui il problema di Newcomb genera fazioni contrapposte ed irriducibili è che la sua formulazione (e con ciò intendo qualsiasi formulazione ne rispetti i termini essenziali) è disseminata di trappole sofistiche. In realtà esiste un concorso di condizioni che determina una forte ambiguità del contesto di scelta; la loro individuazione può aiutare a capire perché si può continuare a discutere del problema senza mai venirne a capo, e perché la discussione può avere degli esiti imprevisti, come quelli che si verificano nel caso del 100% di successi da parte del previsore.

Trappola n° 1: l'ineffabile previsore

In tutte le presentazioni del problema di Newcomb compare la descrizione del previsore come di un individuo dotato di poteri straordinari. Tale descrizione lascia intendere che i poteri in questione siano una caratteristica intrinseca del previsore. Questo modo di porre il problema è ciò che dà origine all'argomento logico usato dai sostenitori della tesi a favore della seconda scatola. L'argomento logico fa leva sulla definizione di "previsore perfetto": qualsiasi cosa uno scelga, il previsore la prevederà. Questa caratteristica potrebbe rientrare anche in un'ottima definizione a priori della natura di Dio. Ma una cosa è trovare una buona definizione della natura di Dio e trarne le dovute conseguenze logiche, un'altra cosa è dimostrare che Dio esiste, e un'altra cosa ancora è dimostrare che un particolare individuo è Dio. Si può certamente dimostrare che è intrinseco alla natura del previsore perfetto (o quasi perfetto) prevedere perfettamente (o quasi perfettamente), ma questo non ha niente a che fare con quello che succede nel mondo reale. Nessuna statistica, neppure una statistica che registri il 100% di successi, può darci una reale informazione sulla necessità logica cui fanno riferimento i sostenitori della tesi a favore della scelta di una sola scatola. Si può definire a priori il concetto di uomo in modo tale da rendere intrinseco alla sua natura il fatto di essere mortale. In questo caso "tutti gli uomini sono mortali" risulterebbe una verità logicamente necessaria, e sarebbe impossibile dimostrare che alcuni uomini non sono mortali. Tuttavia, questo non avrebbe niente a che fare con le statistiche sulla mortalità degli uomini, benché fino a questo momento tali statistiche sembrino avere raggiunto una misura del 100%. Sulla base di tali statistiche, siamo ovviamente in grado di fare previsioni sulla mortalità di un

individuo che presenti caratteristiche tali da essere classificato come "uomo", ma non siamo affatto in grado di stabilire come verità logicamente necessaria il fatto che tale individuo è destinato a morire. Se l'individuo in questione si dimostrasse immortale e volessimo conservare la nostra definizione a priori, saremmo obbligati a rivedere la nostra classificazione, e ad eliminarlo dal genere "uomo". La stessa cosa succede per il previsore. Il fatto che egli abbia previsto con grande successo in tutti i casi verificatisi fino a questo momento ci consente di fare previsioni sui suoi successi futuri, ma non ci consente in nessun modo di stabilire che il prevedere è intrinseco alla sua natura. L'argomento logico, quindi, ha validità solo nel mondo della logica, e in nessun modo può essere applicato al problema di Newcomb, che si presenta come un problema del mondo reale.

Perché i sostenitori della tesi per cui si deve scegliere una sola scatola non tengono conto di queste ovvie considerazioni? Credo che ciò dipenda dalla definizione del problema. E' ovvio che, una volta che si è definito un particolare individuo come un essere la cui intrinseca natura è quella di prevedere perfettamente, diventa una verità necessaria il fatto che sia impossibile immaginare una situazione in cui non c'è connessione fra le scelte dei soggetti sottoposti all'esperimento e le sue decisioni. La necessità logica discende dalla definizione di previsore che è fornita dal problema stesso. Tale definizione gioca evidentemente sull'ambiguità dei piani del discorso: da una parte, poiché il problema è evidentemente immaginario, si suppone che sia inevitabile fornire una definizione a priori delle sue condizioni; dall'altra, poiché si vuole che tale problema immaginario sia concepito come un problema reale, le condizioni epistemologiche di accettabilità delle credenze risultano essere le stesse del mondo reale, cioè le condizioni che escludono che una definizione a priori possa produrre conseguenze fattuali. Ma in questo modo si fa confusione, dal momento che la libertà di definizione del problema che è lasciata al suo inventore non è trasferibile all'interno del problema stesso. Si può essere disposti ad accettare che la definizione del problema di Newcomb offra una descrizione a priori, e presenti fra i dati un previsore intrinsecamente perfetto; però, se si accetta di fare l'esperimento mentale di trasferirsi all'interno del problema, non si possono conservare queste informazioni. In breve: se immagino di essere uno degli individui sottoposti all'esperimento, cui viene spiegato il meccanismo sul luogo stesso del fatto, non posso anche immaginare di riuscire a credere a qualcuno che mi parli di un previsore intrinsecamente perfetto. Per poter sapere che un qualche individuo particolare è un previsore di questo genere, la persona che mi informa di questo fatto dovrebbe essere un previsore perfetto. E per poter credere a questo fatto, anch'io dovrei esserlo. A meno di trovarmi di fronte ad evidenti violazioni delle leggi di natura, cioè ad un miracolo. E con questo veniamo alla seconda trappola.

Trappola n° 2: l'impossibilità di non credere

La maggior parte dei filosofi che hanno discusso il problema di Newcomb ha insistito sul fatto che fra le sue condizioni c'è anche quella per cui il soggetto sottoposto al problema ha un'enorme fiducia nelle capacità del previsore, fiducia che è confermata da dati empirici. Questo dovrebbe dare l'illusione che l'argomento a priori non abbia alcuna rilevanza, ma in realtà si tratta solo di un modo diverso di porlo. Non si dice quali siano le *prove* della straordinaria capacità del previsore, ma si dice che il previsore è tanto perfetto da poter superare un numero indefinito di prove. Questo è un modo per ribadire che crediamo, siamo costretti a credere, all'esistenza di un previsore di questo genere, perché, quanto più siamo scettici, tanto maggiori saranno le prove che ci verranno concesse. Tuttavia, io credo che sia profondamente scorretto porre tra le condizioni di un problema teorico quella per cui il soggetto che deve operare la scelta crede fermamente a qualcosa, almeno quando il problema teorico ha soluzioni diverse a seconda di ciò che si crede. I sostenitori della tesi a favore di una sola scatola hanno il solo torto di accettare questa assurda condizione. Il fatto che essi credano ad una cosa priva di senso non è la conseguenza di un qualche loro errore teorico, ma è la conseguenza delle credenze che sono state attribuite loro dal problema stesso. Tali credenze sono non solo assurde, ma anche ingiustificate dal contesto del problema. La descrizione che ci viene offerta sembra presupporre che, per credere alla presenza di capacità straordinarie nel previsore, sia sufficiente un certo numero di previsioni azzeccate. In realtà, però, quello di cui si avrebbe bisogno sarebbe una previsione fatta in condizioni tali da renderla impossibile per una persona normale. Ma questo è esattamente ciò che viene categoricamente escluso dalle condizioni del problema (non è concesso a chi sceglie, per esempio, affidarsi a meccanismi casuali). Allora non si può dire che i soggetti abbiano un'"enorme fiducia" nelle capacità del previsore. O si dice che i soggetti hanno una vera e propria *fede* in tali capacità, o si dice che essi hanno una fiducia moderata, e che le prove a favore di tale fiducia non hanno nulla di straordinario. Altrimenti si dice che le prestazioni del previsore sono talmente fuori dal comune da determinare una fiducia straordinaria. Se fosse presentato nel primo modo, il problema diventerebbe: "immagina di essere un perfetto imbecille; come ti comporteresti di fronte a due scatole ecc...". Se fosse presentato nel secondo modo, il problema condurrebbe indubbiamente alla scelta delle due scatole. Se infine fosse presentato nel terzo modo, il problema condurrebbe inevitabilmente alla scelta di una sola scatola.

Trappola n° 3: la definizione a priori e la statistica

Il fatto che vengano fornite contemporaneamente una definizione a priori e una statistica determina un'ulteriore conseguenza. Questa descrizione induce a credere che la statistica sia la base per l'accettazione della definizione a priori,

cosa che, come ho mostrato a proposito dell'argomento di tipo logico, non ha alcun senso. E' questa la fonte dell'errore che induce alcuni sostenitori della tesi delle due scatole a credere di poter cambiare idea in presenza di una misura statistica del 100%. L'errore consiste nel supporre che una percentuale di successi del 100% faccia scattare una necessità di tipo logico, mentre tale percentuale è semplicemente una registrazione di ciò che è successo, e non può avere per sua natura il carattere definitivo e necessario che è richiesto dalla verità logica. L'alta percentuale di successi e la necessità di tipo logico non vanno infatti necessariamente di pari passo: si potrebbe immaginare ad esempio un mondo in cui esista veramente un previsore "perfetto", nel senso di un previsore che sia realmente in grado di sapere in precedenza ciò che succederà in un dato istante, qualsiasi altra cosa succeda nel frattempo, senza che questo implichi l'assoluta infallibilità di tale previsore. D'altra parte, si può immaginare qualcuno che, per fortuna o per esperienza, abbia indovinato finora cento volte su cento, senza che per questo possa essere descritto come il previsore "perfetto". Quindi, in presenza di una percentuale piena di successi del previsore, non avremmo alcuna ragione di credere in una relazione necessaria fra le sue previsioni e i fatti previsti, così come non avremmo sufficienti ragioni per negare l'esistenza di una tale relazione solo per il fatto di avere registrato degli insuccessi da parte del previsore.

In realtà, credo che la falsa credenza nella possibilità di passare dalla verità probabilistica alla necessità di tipo logico non sia causata solo dall'incongruo accostamento dei due dati nella presentazione del problema. Assieme a questo tranello concorrono anche gli altri due, quello che pone fra le condizioni del problema la fiducia nel previsore e quello che porta ad una confusione fra i vari piani del discorso. Per definizione non può succedere durante l'esperimento qualcosa che falsifichi le credenze dei soggetti. Chi argomenta a favore della tesi dell'unica scatola accetta semplicemente questa condizione: coloro che scelgono una sola scatola decidono in questo modo perché non possono fare una scelta che contraddica le condizioni del problema. Di solito nella realtà non ci si imbatte in questo genere di ostacoli: generalmente, la gente agisce in base alle proprie conoscenze o alle proprie credenze, senza riguardo al fatto che le azioni possano contraddire qualche descrizione a priori del mondo. Scegliere le due scatole nel problema di Newcomb equivale a contraddire una definizione a priori, cioè quella del previsore perfetto.

Tre problemi di facile soluzione

Il problema non presenterebbe più alcun aspetto paradossale se fossero definite precisamente le sue condizioni. A seconda del genere di contesto induttivo delineato, si potrebbero configurare in questo caso tre diversi problemi,

ai quali si potrebbero applicare i tre principali argomenti considerati in queste pagine (quello logico e quello statistico a favore di una scatola sola, e quello a favore delle due scatole):

- nel primo problema l'unica serie di dati a disposizione sarebbe quella relativa alla co-variazione statistica tra le scelte e gli esiti. In questo scenario le persone si troverebbero semplicemente di fronte alle due scatole e ai risultati delle scelte precedenti di altre persone. Di fronte ad uno scenario del genere, le persone tenderebbero a formulare una massima di questo tipo: "chi sceglie due scatole, generalmente porta via un milione, mentre chi sceglie una scatola sola guadagna un miliardo". Non avendo nessun altro dato a disposizione, si supporrebbe l'esistenza di un qualsiasi meccanismo che determini questa relazione, e la scelta più ragionevole sarebbe quella di una sola scatola. Questo è lo scenario appropriato all'argomento di tipo statistico a favore di questa scelta.

- nel secondo problema ai dati appena descritti se ne aggiungerebbero altri. Qualcuno spiegherebbe agli sceglitori la co-variazione statistica raccontando loro del previsore quasi perfetto (potrebbe essere lo stesso previsore a farlo), e aggiungerebbe il fatto che i soldi sono stati messi nelle scatole in un momento precedente, e che non c'è nessun modo di spostarli. Quest'ultimo dato non potrebbe essere provato direttamente. Ma supponiamo che gli sceglitori siano disposti a fidarsi della parola di numerosi testimoni, che hanno assistito all'esperimento dall'inizio, quando il previsore ha messo i soldi nelle scatole. In questo caso, lo sceglitore avrebbe un'idea più precisa del "meccanismo" che determina la relazione fra la scelta e l'esito, e l'argomento statistico non avrebbe più senso. Tuttavia, prima di scegliere, le persone sottoposte alla prova non potrebbero fare a meno di chiedere alla persona che ha spiegato loro quali sono i termini del problema che cosa intenda per "previsore quasi perfetto". Se la risposta fosse che per "previsore quasi perfetto" si intende semplicemente qualcuno che ha una grande esperienza della natura umana, e che conosce i soggetti dell'esperimento, per cui è in grado di fare previsioni più accurate della media delle persone che non si dedicano a questo genere di studi sugli amici, la scelta giusta sarebbe ovviamente quella delle due scatole. Se le persone fossero avvertite in precedenza dell'esperimento, potrebbero tentare di ingannare il previsore simulando i comportamenti tipici da "sceglitore di una sola scatola", per prendere poi all'ultimo momento, a sorpresa, entrambe le scatole (sperando che nel frattempo la lunga consuetudine con i comportamenti simulati non arrivi a trasformare la loro intima natura di "sceglitori proditori" di entrambe le scatole). Tuttavia questo problema non si porrebbe, perché in questo scenario, come negli altri, si suppone che le persone vengano colte alla sprovvista dagli organizzatori dell'esperimento.

La risposta alla domanda sulla natura del previsore potrebbe anche essere un'altra. Le persone potrebbero sentirsi spiegate che il previsore quasi perfetto è

un essere in grado di prevedere che cosa succederà, qualsiasi cosa succeda. In questo caso sarebbe più che legittimo non dare alcun credito a questo genere di affermazioni. Per poter credere ad un'affermazione di questo genere bisognerebbe infatti essere il previsore perfetto, cioè sapere già che cosa succederà e che cosa ha scelto il previsore, e conoscere queste cose *ab aeterno*, assieme a tutte le relazioni fra gli oggetti per qualsiasi descrizione della realtà possibile. A meno che non sia possibile ottenere delle prove di questo fatto, il che trasferirebbe il problema nel terzo scenario:

- in questo caso il previsore avrebbe fornito prove spettacolari della propria straordinaria dote. L'unico modo di farlo sembra essere la previsione di fatti futuri che siano l'esito di un processo casuale (i singoli lanci di una moneta, per esempio). Non è comunque affatto pacifico che prove di questo genere, o qualsiasi altro tipo di prova immaginabile, sarebbero mai sufficienti a generare una simile credenza. Se comunque ci fosse qualcuno che arrivasse a credere nel miracolo della previsione quasi perfetta, lo scenario di questa persona risulterebbe completamente stravolto: un mondo in cui le azioni potessero influenzare il passato, attraverso una relazione di *backward causation*, sarebbe un mondo completamente determinato, in cui le relazioni sarebbero stabilite *ab aeterno*: in questo mondo il tempo non avrebbe più alcuna rilevanza, e non sussisterebbe allora più nessun ostacolo all'argomento di tipo logico.

Anche in presenza di fatti straordinari, alcune persone rimarrebbero scettiche riguardo alle capacità dell'individuo che si presenta come previsore quasi perfetto. Tuttavia, anche in questo caso le prove di tipo statistico cambierebbero completamente significato. In questo caso le persone potrebbero legittimamente dubitare di tutta la storia, e usare le statistiche come prove del fatto che qualcuno ha creato un meccanismo per cui le persone che scelgono una scatola sola prendono quasi sicuramente il miliardo, mentre le persone che scelgono entrambe le scatole guadagnano solo un milione. Rimarrebbe vero che, sia che le persone sottoposte all'esperimento arrivino a credere in un essere in grado di fare miracoli, sia che le persone siano convinte di essere in presenza di un colossale imbroglio, la scelta giusta in questo scenario sarebbe quella di una sola scatola.

Conclusioni

Nella descrizione classica del problema di Newcomb i tre scenari sono fusi assieme. Non esiste alcuna enumerazione precisa dei dati che sono a disposizione di chi deve scegliere. La situazione è definita in modo tale da poter adattarsi a qualsiasi credenza a proposito del previsore e del significato delle statistiche. Nonostante il fatto che gli sceglitori delle due scatole abbiano dalla

loro parte la ragionevolezza e la minore complessità teorica di uno scenario in cui non si verificano fatti miracolosi e non agisce la *backward causation*, la descrizione del problema, in sé, non nega alcuna plausibilità alla tesi avversaria. Resta comunque escluso che le differenze d'opinione riguardino i principi di scelta, se non indirettamente, cioè in relazione ad assunzioni di tipo metafisico. I problemi teorici cui vanno incontro i sostenitori della *backward causation* sono sicuramente degni di considerazione, ma non sono una prerogativa del problema di Newcomb. Neppure le questioni sulla ragionevolezza del comportamento cooperativo (che potrebbero essere chiamate in causa qualora si volesse intendere la descrizione del problema di Newcomb come uno scenario ipotetico, nel quale individui altrettanto ipotetici e personalmente disinteressati dovessero scegliere principi universali di condotta) costituiscono un esito peculiare e caratteristico delle discussioni su questo problema. Tale paradosso continuerà sicuramente ad essere un ottimo pretesto per seguire a discutere del libero arbitrio, del determinismo, e di molti problemi epistemologici fondamentali; ma di per sé non presenta nessuna difficoltà che possa dare da pensare ad un filosofo, e tantomeno ad una persona che si trovasse realmente alle prese con una scelta di questo genere.

Per ulteriori approfondimenti e per i riferimenti bibliografici cfr. V. Ottonelli, 'Le trappole di Newcomb', *Rivista di filosofia*, LXXXIV, 1, 1993, pp. 49-71.